Stoch Environ Res Risk Assess
DOI 10.1007/s00477-007-0178-2

ORIGINAL PAPER

Conditional simulations of water–oil flow in heterogeneous porous media

Mingjie Chen · Zhiming Lu · George A. Zyvoloski

Abstract This study is an extension of the stochastic analysis of transient two-phase flow in randomly heterogeneous porous media (Chen et al. in Water Resour Res 42:W03425, 2006), by incorporating direct measurements of the random soil properties. The log-transformed intrinsic permeability, soil pore size distribution parameter, and van Genuchten fitting parameter are treated as stochastic variables that are normally distributed with a separable exponential covariance model. These three random variables conditioned on given measurements are decomposed via Karhunen–Loève decomposition. Combined with the conditional eigenvalues and eigenfunctions of random variables, we conduct a series of numerical simulations using stochastic transient water–oil flow model (Chen et al. in Water Resour Res 42:W03425, 2006) based on the KLME approach to investigate how the number and location of measurement points, different random soil properties, as well as the correlation length of the random soil properties, affect the stochastic behavior of water and oil flow in heterogeneous porous media.

Keywords Karhunen–Loève decomposition · Conditional simulation · Water–oil flow · Uncertainty · NAPL

1 Introduction

Many hazardous organic materials, such as oils, gasoline, and petrochemicals are widely used in the chemical and petroleum industry. Accidental releases of these nonaqueous phase liquids (NAPL) to the subsurface are inevitable and represent a significant threat to water supply and ecosystems. Although the solubility of NAPLs is often low in water, the concentration often exceeds drinking water standards. Thus, a small amount of NAPL can contaminate large volumes of groundwater over a long period of time. Therefore, it is very important to understand the processes associated with contaminant migration and fate. Numerical multiphase flow models are used to study the various aspects of these processes in order to conduct risk assessment and design of cost-efficient remediation (e.g. Abriola 1989). Another primary application of multiple fluid systems is petroleum reservoir engineering, which depends on the understanding of reservoir mechanics to design schemes for efficient oil recovery. A petroleum reservoir is a complicated mixture of hydrocarbon fluids, brine, porous rock and fractures. The structure of the void space is tortuous and heterogeneous, which influence and even dominate the fluid flow. Complete characterization of subsurface properties is impossible. Many researchers have resorted to stochastic modeling of subsurface flow in the last two decades (Dagan 1989; Gelhar 1993; Zhang 2002). This approach assumes hydrological variables are uncertain and treats them as stochastic processes; this defines flow model in a stochastic context rather than in the traditional deterministic framework, and predicts the output in terms of moments, usually the means and covariances.

Zhang and Lu (2004) proposed a new stochastic approach, called KLME, based on Karhunen–Loève decomposition and polynomial expansions, and applied it to saturated flow models. Yang et al. (2004) extended KLME to analysis of saturated–unsaturated flow described by Richard's equation. Chen et al. (2005, 2006) developed a stochastic multiphase flow model following the

M. Chen (✉) · Z. Lu · G. A. Zyvoloski
Hydrology & Geochemistry Group (EES-6),
Los Alamos National Laboratory, Los Alamos, NM, USA
e-mail: mchen@lanl.gov

same approach. These studies demonstrated the accuracy and efficiency of KLME over traditional Monte Carlo simulation or other stochastic moment approaches. However, the accuracy of KLME method, compared to Monte Carlo results, decrease as the variations of soil properties increase. Zhang and Lu (2004) show the KLME method is very accurate when the variance of the log hydraulic conductivity is 2.0, and not good at 4.0. Many studies in conditional simulation indicate that conditioning on measurements of the log hydraulic conductivity can reduce its overall uncertainty, especially in the vicinity of the conditioning points, which may reduce the predictive uncertainties of flow and transport (Dagan 1982; Guadagnini and Neuman 1999a, b; Lu et al. 2002; Tartakovsky et al. 1999). Therefore, conditional simulation will enable KLME to be applied effectively in flow systems with strongly heterogeneous fields and more accurate in moderate heterogeneous soil. Lu and Zhang (2004) conducted conditional simulations of saturated flow using KLME by incorporating measurements of the log hydraulic conductivity. The key step is to derive and calculate the conditional eigenvalues and eigenfuctions through unconditional ones of log hydraulic conductivity using kriging techniques. Running the stochastic saturated flow model with these conditional eigenvalues and eigenfuctions leads to the conditional simulations. In this study, we extend their work from the saturated flow to the transient water–oil flow system with three random input variables, including log intrinsic permeability Y, log pore size distribution parameter β, and log van Genuchten fitting parameter $\bar{n}$, with measurements in various locations. We design a series of scenarios simulated by the stochastic KLME flow model to examine how the location and number of measurements, and the correlation length of covariance of these three random input variables, influence the magnitude and distribution of uncertainties of predictive variables.

2 KL decomposition of conditional random field

A short description of Karhunen–Loève (KL) decomposition of unconditional random fields is given below for readers to better understand conditional cases. The KL decomposition of a stochastic process $\alpha(\mathbf{x},\theta)$ is based on the spectral decomposition of the covariance function of α, $C_\alpha(\mathbf{x},\mathbf{y})$, with a set of orthogonal polynomials. Here, $\mathbf{x}$ and $\mathbf{y}$ are spatial locations, and the argument θ denotes the random nature of the corresponding quantity. $C_\alpha(\mathbf{x},\mathbf{y})$ is symmetrical and positive definite, whose eigenfunctions are mutually orthogonal and form a complete set spanning the function space to which $\alpha(\mathbf{x},\theta)$ belongs (Ghanem and Dham 1998). The mean-removed stochastic process $\alpha'(\mathbf{x},\theta)$ can be expanded as follows (Karhunen 1947; Loève 1948; Zhang and Lu 2004; Chen et al. 2005):

$$\alpha'(\mathbf{x},\theta) = \sum_{n=1}^{\infty} \xi_n(\theta)\sqrt{\lambda_n}\phi_n(\mathbf{x}), \tag{1}$$

where, λ_n and $\phi_n(\mathbf{x})$ are the eigenvalues and eigenfunctions of the covariance kernel $C_\alpha(\mathbf{x},\mathbf{y})$, respectively. $\phi_n(\mathbf{x})$ are orthogonal deterministic functions and form a complete set

$$\int_\Omega \phi_n(\mathbf{x})\phi_m(\mathbf{x})d\mathbf{x} = \delta_{nm}, \tag{2}$$

where Ω denotes the spatial domain where $\alpha(\mathbf{x},\theta)$ is defined. Eigenvalues and eigenfunctions can be solved from the Fredholm equation

$$\int_\Omega C_\alpha(\mathbf{x},\mathbf{y})\phi(\mathbf{x})d\mathbf{x} = \lambda\phi(\mathbf{y}). \tag{3}$$

As defined, $\{\xi_n(\theta)\}$ forms a set of orthogonal random variables, and has properties of $\langle\xi_n(\theta)\rangle = 0$, and $\langle\xi_n(\theta)\xi_m(\theta)\rangle = \delta_{nm}$, and δ_{nm} is the Krönecker delta function. For separable exponential covariance function such as

$$C_\alpha(\mathbf{x},\mathbf{y}) = \sigma_\alpha^2 \exp\left(-\sum_i \frac{|x_i - y_i|}{\eta_i}\right), \tag{4}$$

where σ_α^2 and η_i are the variance and correlation length of $\alpha(\mathbf{x},\theta)$ in the ith direction, analytical solutions of eigenvalues and eigenfunctions can be obtained by combing one dimensional analytical solution in each direction. For the general case, the eigenvalues and eigenfunctions have to be solved numerically via Galerkin-type method (Ghanem and Spanos 1991).

Assume α is measured at N_α locations, we can obtain the conditional mean $\langle\alpha\rangle^{(C)}$ and variance $C_\alpha^{(C)}$ using the kriging method:

$$\langle\alpha(\mathbf{x})\rangle^{(C)} = \langle\alpha(\mathbf{x})\rangle + \sum_{i=1}^{N_\alpha} \mu_i(\mathbf{x})[\alpha(\mathbf{x}_i) - \langle\alpha(\mathbf{x}_i)\rangle], \tag{5}$$

$$C_\alpha^{(C)}(\mathbf{x},\mathbf{y}) = C_\alpha(\mathbf{x},\mathbf{y}) - \sum_{i,j=1}^{N_\alpha} \mu_i(\mathbf{x})\mu_j(\mathbf{y})C_\alpha(\mathbf{x}_i,\mathbf{x}_j), \tag{6}$$

where $\langle\alpha(\mathbf{x})\rangle$ and $C_\alpha(\mathbf{x},\mathbf{y})$ are unconditional mean and covariance. Weighting coefficients $\mu_i(\mathbf{x})$ represent the relative significance of each measurement $\alpha(\mathbf{x}_i)$ in predicting the $\langle\alpha\rangle^{(C)}$ at location $\mathbf{x}$, and can be determined from the following equations:

$$\sum_{i=1}^{N\alpha} \mu_i(\mathbf{x}) C_\alpha(\mathbf{x}_i, \mathbf{x}_j) = C_\alpha(\mathbf{x}, \mathbf{x}_j), \quad j = 1, 2, \ldots, N_\alpha. \tag{7}$$

Apparently, the two-point conditional covariance $C_\alpha^{(C)}(\mathbf{x},\mathbf{y})$ is no longer stationary as the unconditional covariance, and depends on the locations $\mathbf{x}$ and $\mathbf{y}$, instead of their separation distance. Generally, the corresponding eigenvalues $\lambda_n^{(C)}$ and eigenfunctions $\phi_n^{(C)}$ have to be solved numerically (Ghanem and Spanos 1991), which is comparably expensive, especially if one needs to update eigenvalues and eigenfunctions sequentially because of introducing new conditioning points. Lu and Zhang (2004) related the conditional eigenvalues and eigenfuctions to their unconditional counterpart λ_n and ϕ_n. The later can be solved easily for a special case such as a two-dimensional rectangular or three-dimensional brick-shaped domain. Since the unconditional eigenfuctions ϕ_n are complete, the basic idea of this algorithm is to expand $\mu_i(\mathbf{x})$ and $\phi_n^{(C)}$ based on ϕ_n (Lu and Zhang 2004):

$$\mu_i(\mathbf{x}) = \sum_{k=1}^{\infty} \mu_{ik} \phi_k(\mathbf{x}), \tag{8}$$

$$\phi^{(C)}(\mathbf{x}) = \sum_{k=1}^{\infty} d_k \phi_k(\mathbf{x}), \tag{9}$$

where μ_{ik} is coefficients determined by

$$\sum_{i=1}^{N_\alpha} C_\alpha(\mathbf{x}_i, \mathbf{x}_j)\mu_{ik} = \lambda_k \phi_k(\mathbf{x}_j), \quad j = 1, 2, \ldots, N_\alpha, \quad k = 1, 2, \ldots \tag{10}$$

and d_k is components of eigenvector solved from the following eigen-problem.

$$\lambda_k d_k - \sum_{m=1}^{\infty} \left(\sum_{i,j=1}^{N\alpha} C_\alpha(\mathbf{x}_i, \mathbf{x}_j)\mu_{im}\mu_{ik} \right) d_m = \lambda^{(C)} d_k, \quad k = 1, 2, \ldots \tag{11}$$

The computational efforts of obtaining conditional eigenvalues and eigenfunctions in this way are much less than that of solving them numerically via Galerkin-type methods. Moreover, the conditional eigenvalues and eigenfuctions can be updated from the existing ones when more measurements are available. This kind of update costs less computing and hence is more efficient than a new calculation.

3 KL-based conditional moment equations

The following governing equations are derived from the conservation equations and Darcy's relationship for the transient water–oil phase flow (Abriola and Pinder 1985):

$$\frac{\partial^2 P_l(\mathbf{x},t)}{\partial x_i^2} + \frac{\partial Z_l(\mathbf{x},t)}{\partial x_i}\left[\frac{\partial P_l(\mathbf{x},t)}{\partial x_i} + \rho_l g \delta_{i1}\right] = \exp\left[-Z_l(\mathbf{x},t)\right]\left[\phi \frac{\partial S_l(\mathbf{x},t)}{\partial t} - F_l(\mathbf{x},t)\right], \tag{12}$$

where $P_l(\mathbf{x},t)$ is the fluid pressure; l denotes liquids ($l = w, o$); $S_l(\mathbf{x},t)$ are the water ($l = w$) and oil ($l = o$) saturations; $\mathbf{x}$ is the position vector in 2- or 3-D; $F_l(\mathbf{x},t)$ is a source or sink term; $Z_l(\mathbf{x},t) = \ln \lambda_l(\mathbf{x},t)$, and $\lambda_l(\mathbf{x},t) = k(\mathbf{x})k_{rl}(S_l)/\mu_l$ is liquid mobility, where $k(\mathbf{x})$ is the intrinsic permeability of porous media, k_{rl} is the water or oil relative permeability, and μ_l is the liquid dynamic viscosity; ρ_l is fluid density, and ϕ is the porosity of the media. δ_{i1} is the Krönecker delta function, which equals 1 when i is 1 (upward direction) or 0 otherwise. The initial and boundary conditions are as follows:

$$P_l(\mathbf{x}, 0) = P_{l0}(\mathbf{x}), \quad \mathbf{x} \in \Omega, \tag{13}$$

$$P_l(\mathbf{x}, t) = P_{lt}(\mathbf{x}, t), \quad \mathbf{x} \in \Gamma_D, \tag{14}$$

$$\mathbf{n}_i(\mathbf{x}) \exp\left[Z_l(\mathbf{x},t)\right]\left[\frac{\partial P_l(\mathbf{x},t)}{\partial x_i} + \rho_l g \delta_{i1}\right] = -Q_l(\mathbf{x},t), \mathbf{x} \in \Gamma_N, \tag{15}$$

where $P_{l0}(\mathbf{x})$ is the initial pressure in the domain Ω ; $P_{lt}(\mathbf{x},t)$ is the prescribed pressure on Dirichlet boundary segments Γ_D ; $Q_l(\mathbf{x},t)$ is the prescribed fluid flux across Neumann boundary segments Γ_N; g is the gravity; $\boldsymbol{n}(\mathbf{x})$ is the outward unit vector normal to the boundary Γ_N.

van Genuchten (1980) model is used to describe the relationship between saturation, capillary pressure, and relative permeability:

$$k_{rw} = \bar{S}_w^{1/2}\left[1 - \left(1 - \bar{S}_w{}^{1/m}\right)^m\right]^2, \tag{16}$$

$$k_{ro} = (1 - \bar{S}_w)^{1/2}\left(1 - \bar{S}_w{}^{1/m}\right)^{2m}, \tag{17}$$

$$\overline{S}_w = \left[1 + (\alpha P_c)^n\right]^{-m}, \tag{18}$$

where $\overline{S}_w = (S_w - S_{wr})/(1 - S_{wr})$ is the effective water saturation, $S_w = 1 - S_o$ is water saturation, and S_{wr} is the residual water saturation, α is the pore size distribution parameter, $P_c = P_o - P_w$ is the capillary pressure, n is the van Genuchten fitting parameters and $m = 1 - 1/n$. In our study, the log intrinsic permeability, log pore size distribution parameter, and log van Genuchten fitting parameter $\bar{n} = \ln(n - 1)$ are considered random fields, and expanded using the conditional KL decomposition described in the last section. The governing Eqs. (12)–(15) are the stochastic water–oil flow model, and are solved via KL-based perturbation methods.

The mathematical formulation of the Eqs. (12)–(15) using the KLME method are presented by Chen et al. (2006) in detail. The dependent variables, such as water pressure P_w, are written as $P_w = P_w^{(0)} + P_w^{(1)} + \ldots$, where $P_w^{(n)}$ is the nth order term in terms of standard deviation of soil properties. These nth order terms can be decomposed in terms of a set of orthogonal Gaussian random variables $\{\xi_n\}$ and deterministic coefficients. For example, $P_w^{(1)} = \sum_{j=1}^{\infty} \xi_j P_{w,j}^{(1)}$, where $P_{w,j}^{(1)}$ are deterministic coefficients. Substituting these equations to the original stochastic governing equations yields a series of deterministic governing equations as shown in Eqs. (19)–(22). In summary, the idea of the KLME approach is to decompose stochastic governing equations of flow into a series of deterministic equations, which can be solved using existing numerical simulators. The solutions are then assembled to obtain explicit moments (means and variances) of the dependent variables. The KL-based conditional moment equations can be derived in the similar way. Zeroth order equations are shown as

$$\frac{\partial^2 P_w^{(0)}}{\partial x_i^2} + \frac{\partial Z_w^{(0)}}{\partial x_i}\left[\frac{\partial P_w^{(0)}}{\partial x_i} + \rho_w g \delta_{i1}\right] = \frac{C_{ow}^{(0)}}{e^{Z_w^{(0)}}}\frac{\partial P_c^{(0)}}{\partial t} - \frac{F_w}{e^{Z_w^{(0)}}}, \tag{19}$$

$$\frac{\partial^2 P_o^{(0)}}{\partial x_i^2} + \frac{\partial Z_o^{(0)}}{\partial x_i}\left[\frac{\partial P_o^{(0)}}{\partial x_i} + \rho_o g \delta_{i1}\right] = \frac{-C_{ow}^{(0)}}{e^{Z_w^{(0)}}}\frac{\partial P_c^{(0)}}{\partial t} - \frac{F_o}{e^{Z_w^{(0)}}}, \tag{20}$$

and first order equations are shown as

$$\begin{aligned}&\frac{\partial^2 P_{w,j}^{(1)}}{\partial x_i^2} + J_{wi}\frac{\partial Z_{w,j}^{(1)}}{\partial x_i} + \frac{\partial Z_w^{(0)}}{\partial x_i}\frac{\partial P_{w,j}^{(1)}}{\partial x_i} \\ &= \frac{C_{ow}^{(0)}}{e^{Z_w^{(0)}}}\left[\frac{\partial P_{c,j}^{(1)}}{\partial t} - Z_{w,j}^{(1)}\frac{\partial P_c^{(0)}}{\partial t}\right] + \frac{C_{ow,j}^{(1)}}{e^{Z_w^{(0)}}}\frac{\partial P_c^{(0)}}{\partial t} - \frac{F_w}{e^{Z_w^{(0)}}}Z_{w,j}^{(1)},\ j = \overline{1,M}\end{aligned} \tag{21}$$

$$\begin{aligned}&\frac{\partial^2 P_{o,j}^{(1)}}{\partial x_i^2} + J_{oi}\frac{\partial Z_{o,j}^{(1)}}{\partial x_i} + \frac{\partial Z_o^{(0)}}{\partial x_i}\frac{\partial P_{o,j}^{(1)}}{\partial x_i} \\ &= \frac{-C_{ow}^{(0)}}{e^{Z_o^{(0)}}}\left[\frac{\partial P_{c,j}^{(1)}}{\partial t} - Z_{o,j}^{(1)}\frac{\partial P_c^{(0)}}{\partial t}\right] - \frac{C_{ow,j}^{(1)}}{e^{Z_o^{(0)}}}\frac{\partial P_c^{(0)}}{\partial t} - \frac{F_o}{e^{Z_o^{(0)}}}Z_{o,j}^{(1)},\ j = \overline{1,M}\end{aligned} \tag{22}$$

where $C_{ow}(\mathbf{x},t) = \phi(1 - S_{wr})\frac{\partial \bar{S}_w}{\partial P_c}$ is a intermediate stochastic variable, accounting for the derivative of effective water saturation with respect to capillary pressure. Z_l are the log phase mobility of l phase ($l = w, o$), F_l are the l phase source/sink terms, and $J_{li} = \partial P_l^{(0)}/\partial x_i + \rho_l g \delta_{i1}$. Please note we omit ($\mathbf{x}$, t) for spatio-temporal variables in the above equations to simplify the expression. There is only one set of zeroth order equations, but M sets of first order equations. M is the number of terms needed to capture the most of uncertainty. In our examples, we choose 200 of M. Solve the series of equations, and we can construct the mean and variance of liquid pressures,

$$\langle P_l(\mathbf{x},t)\rangle \approx \left\langle P_l^{(0)}(\mathbf{x},t)\right\rangle + \left\langle P_l^{(1)}(\mathbf{x},t)\right\rangle = P_l^{(0)}(\mathbf{x},t), \tag{23}$$

$$C_{P_l}(\mathbf{x},\mathbf{y},t) = \sum_{j=1}^{M} P_{l,j}^{(1)}(\mathbf{x},t)P_{l,j}^{(1)}(\mathbf{y},t), l = o, w. \tag{24}$$

The variances of fluid saturation can be found in the similar way. This stochastic KLME water–oil flow model was coded using Fortran.

4 Illustrative examples

In this section, we use the stochastic water-oil flow model to examine how the random input variables and their measurements influence stochastic behavior of fluid pressure and saturation. We didn't conduct Monte-Carlo simulations for these cases in this study, since the accuracy and efficiency of the developed stochastic model have been demonstrated by Chen et al. (2006), and our goal is to use it to investigate the problems of interest.

We assume the log intrinsic permeability $Y(\mathbf{x})$, log pore-size distribution parameter $\beta(\mathbf{x})$, and log van Genuchten fitting parameter $\bar{n}(\mathbf{x})$ to be random fields with separable exponential covariance functions as Eq. (4). We consider a two-dimensional domain in a water–oil flow system in hypothetic heterogeneous porous medium. This vertical cross section is 3 m deep and 0.96 m wide, uniformly discretized into 50×16 square elements of 0.06 m $\times$ 0.06 m. The no-flow conditions are prescribed at two lateral boundaries. The water and oil pressure are specified at the bottom of the domain (Fig. 1). Oil is leaked into the domain at node $X1$ = 2.4 m, $X2$ = 0.48 m (black solid circle in Fig. 1) and a constant precipitation rate is fixed at the top boundary. The soil porosity is 0.5. The initial water saturation around the oil leak point is about 0.98. The initial values and boundary conditions represent a continuous DNAPL leak into the nearly clean soil with constant precipitation on the soil surface.

The primary fluids properties, soil properties, and boundary conditions for the baseline case are listed in Table 1. The measurements data are extracted from the "true" fields, generated from the unconditional Karhunen–Loève decomposition (Eq. 1). Based on this unconditional simulation of baseline case, we will conduct conditional simulations with different number of random variables,

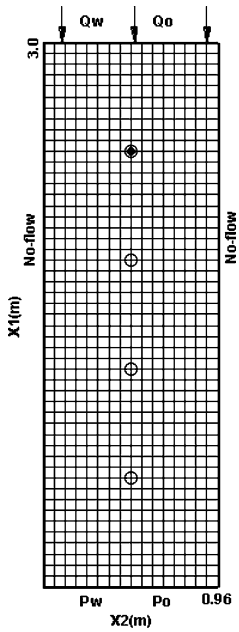

Fig. 1 Model domain

measurements, and correlation length (Table 2). All the cases are simulated to 1 day.

4.1 Number of measurements

Case 1 is the conditional counterpart of the unconditional baseline case, with measurements of $Y(\mathbf{x})$, $\beta(\mathbf{x})$, and $\bar{n}(\mathbf{x})$ at the location of oil leak (0.48, 2.4 m), while Case 2 have additional three measurements at X1 = 1.8, 1.2, 0.6 m along the central vertical line. All the four measurement locations are denoted with small circle in Fig. 1. To simulate conditional Case 1 and Case 2, we first solve for the unconditional eigenvalues and eigenfunctions, then compute conditional ones using the algorithm described in Sect. 2. The first 50 terms of unconditional and conditional eigenvalues of covariance of $\beta(\mathbf{x})$ and their cumulations are shown in Fig. 2, where condition 1 has only one measurement at X1 = 2.4 m and condition 2 has all the four measurements. It is seen that conditional eigenvalues with four measurements is less than those with one measurement, which, in turn, are less than those of the unconditional eigenvalues. It indicates that variability of $\beta(\mathbf{x})$ in Case 2 are smaller than that in both Case 1 and unconditional baseline case. It is shown that the series of eigenvalues is monotonously decreasing and the first 50 terms can account for about 90% of the total variability. Figure 3 compares the third term of unconditional eigenfunctions, conditional eigenfunctions with one measurement (condition 1) and four measurements (condition 2). It is obvious that the measurements affect the characteristic values and scales of eigenfunction fields.

The conditional means and variances of $\beta(\mathbf{x})$ are shown in Figs. 4 and 5, which will be used in Case 1 and Case 2. The values of means around measurements are significantly influenced by the measurements. The variability decreases around the measurements in a radiated distribution, and the overall variability of $\beta(\mathbf{x})$ used in Case 2 is a little bit smaller than that used in Case 1, since Case 2 has all the four measurements honored in the random fields. The other two random variables $Y(\mathbf{x}), \bar{n}(\mathbf{x})$ have the similar behaviors in the two cases.

Figure 6 compares the variances of water saturation simulated from Baseline Case, Case 1, and Case 2. For the unconditional case, the variances of water saturation

Table 1 Soil and fluid properties and boundary conditions

Parameter name	Symbol	Units	Baseline Case
Water density	ρ_w	kg/m^3	997.81
Oil density	ρ_o	kg/m^3	1,500
Water viscosity	μ_w	Pa s	1.0×10^{-3}
Oil viscosity	μ_o	Pa s	6.5×10^{-4}
Mean intrinsic permeability	$<k>$	m^2	3.78×10^{-11}
Mean pore size distribution	$<\alpha>$	1/Pa	1.23×10^{-4}
Mean fitting parameter n	$<n>$	–	1.35
Variance of permeability	σ_k^2	–	5.20×10^{-22}
Variance of pore size distribution	σ_α^2	–	1.55×10^{-10}
Variance of fitting parameter n	σ_n^2	–	1.13×10^{-2}
Coefficient of variation (k)	$CV(k)$	–	53.29%
Coefficient of variation (α)	$CV(\alpha)$	–	10.03%
Coefficient of variation (n)	$CV(n)$	–	7.86%
Correlation length	$\eta_k,\eta_\alpha,\eta_n$	m	0.3
Lower boundary water pressure	P_w	Pa	1.40×10^5
Lower boundary oil pressure	P_o	Pa	1.55×10^5
Upper boundary water flux	Q_w	m/s	1.0×10^{-8}
Oil leakage rate	F_o	kg/day	100

Table 2 Cases simulated in this study

Case	Conditioned variables	Number of measurements	Correlation length (m)	Oil source (kg/day)
Baseline	0	–	0.3	100
Case 1	$Y, \beta, \bar{n}$	1	0.3	100
Case 2	$Y, \beta, \bar{n}$	4	0.3	100
Baseline I	0	–	0.3	10,000
Case 3	$Y, \beta, \bar{n}$	1	0.3	10,000
Case 4	$Y, \beta, \bar{n}$	4	0.3	10,000
Case 5	Y,β	1	0.3	10,000
Case 6	$Y, \bar{n}$	1	0.3	10,000
Case 7	$\beta, \bar{n}$	1	0.3	10,000
Baseline II	0	–	0.15	100
Case 8	$Y, \beta, \bar{n}$	1	0.15	100

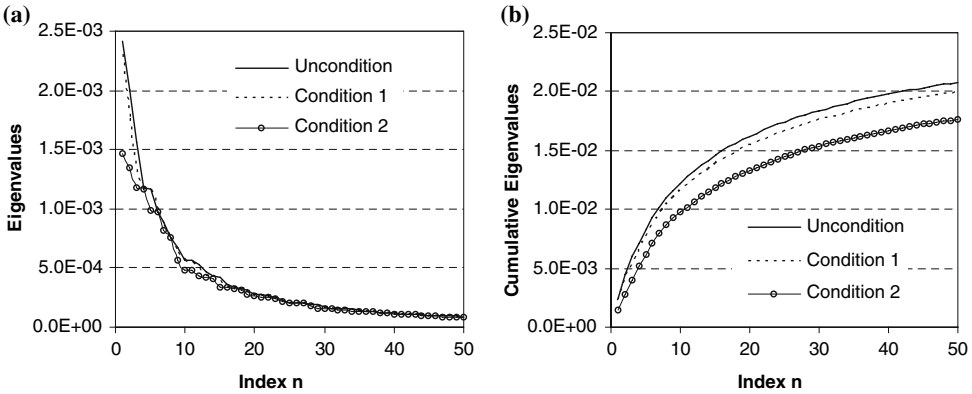

Fig. 2 Unconditional and conditional **a** eigenvalues, **b** cumulative eigenvalues

present a radiated downward distribution around the oil leak location, indicating more uncertainty about the location of the oil as the oil migrates downwards (oil is denser than water). The variances at the bottom are zero since we use constant pressure boundary at the bottom. For the conditional cases (Case 1, 2), the conditioning effect is very localized, reducing the uncertainty of water saturation around the conditioning points. The variances of oil pressure and capillary pressure along the central vertical line of the domain for the three cases are shown in Fig. 7. It is seen that the peak of the profile for Baseline Case at the oil leak location are damped for Case 1 and 2, and the reduction of variability extends to low part of the domain for Case 2, since there are additional three measurements below the measurement at the oil leak location. In addition, the overall oil and capillary pressure variability are reduced for the two conditional cases, compared to the unconditional case. The average variance of oil pressure across the profile is reduced relatively by 92.50% in Case 1, and 94.13% in Case 2 from Baseline Case, respectively. For capillary pressure, the figures are 93.21 and 93.50%, respectively. That means more than 90% of the overall uncertainty of oil or capillary pressure along the vertical central line is reduced. In addition to the number of measurements, the location of measurements is another factor influencing the reduction of uncertainty. As is shown in Fig. 7, the measurement at the oil leak location, where the prediction is most uncertain, contributes much more to the uncertainty reduction than the other three measurements.

We also examine the effect of the oil source term to the prediction uncertainty. We increase the oil injection rate from 100 kg/day in Baseline Case, Case 1 and 2 to 10,000 kg/day in Baseline Case I, Case 3 and 4. The variance profiles of oil pressure and capillary pressure are presented in Fig. 8. The variance profiles from the unconditional Baseline Case I, conditional Case 3 with one measurement, and Case 4 with four measurements show the same behavior as those in the low oil injection rate (Fig. 7), but the magnitude of is almost 4 orders greater. It indicates that the strength of oil source term has the overwhelming effect on the magnitude of prediction variance over the uncertainty of soil properties, from which the

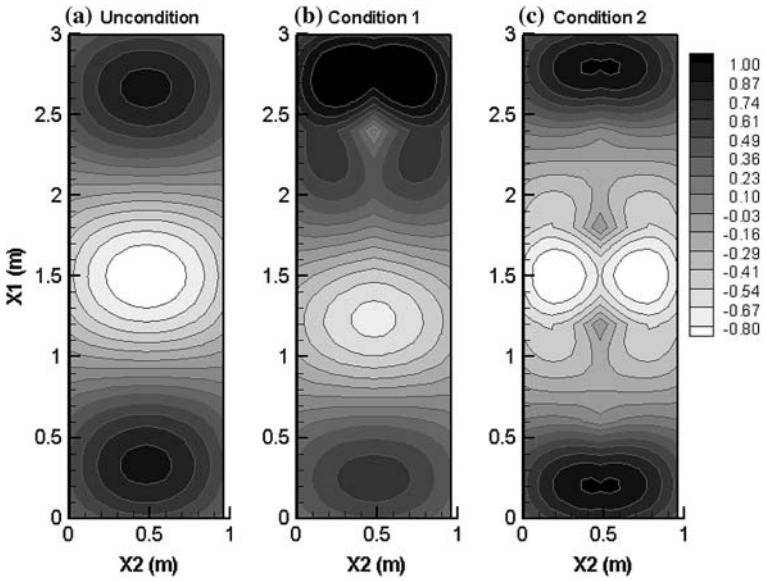

Fig. 3 The third eigenfunctions ($n = 3$) of **a** unconditional, **b** 1 measurement conditional, and **c** 4 measurement conditional simulations

prediction variability originate. The role of the former one in the prediction uncertainty is like amplifier, while the latter is the sound source.

4.2 Number of conditioned random fields

All the three random fields are conditioned simultaneously at one or four locations in Case 1 to 4. In this section, we modify Case 3, and use one of unconditional $\bar{n}, \beta, Y$ field in Case 5, 6, 7 respectively, and leave the other two fields conditional (Table 2). By simulating these cases, we attempt to find the relative importance of the three random variables to the prediction uncertainty.

Figure 9a presents the variance profile of oil pressure along the central vertical line of the modeling domain for Case 3, 5, 6 and 7. With β and $\bar{n}$ conditioned at the oil leak location, variances of oil pressure from Case 7 are smaller than those from unconditional Baseline Case I (Fig. 8), but well above those from Case 3, in which all the three random variables $Y, \beta, \bar{n}$ are conditioned at the oil leak location. Compared to Case 5, in which Y,β are conditioned, and Case 6, in which $Y, \bar{n}$ are conditioned, the reduction of oil pressure variability is the smallest in Case 7. It seems that being conditioned on Y measurements is more efficient in reducing uncertainty of oil pressure than $\beta, \bar{n}$. The difference between conditioned β and $\bar{n}$ is not so large, although, the reduction of oil pressure variability with β conditioned (Case 5) is more than that with $\bar{n}$ conditioned (Case 6).

The variances behavior around the oil leak location (X1 = 2.4 m) presents different characteristic between the cases in Fig. 9a. Case 7 is simulated with β and $\bar{n}$ conditioned at the oil leak location, but the oil pressure variances

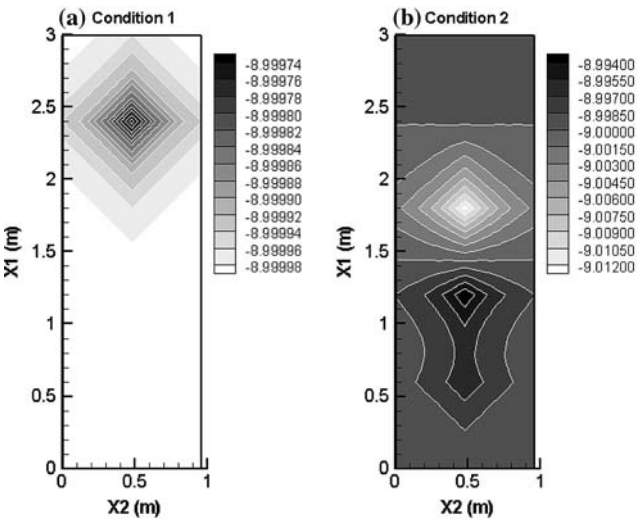

Fig. 4 The means of log pore size distribution β with **a** 1 measurement, and **b** 4 measurements

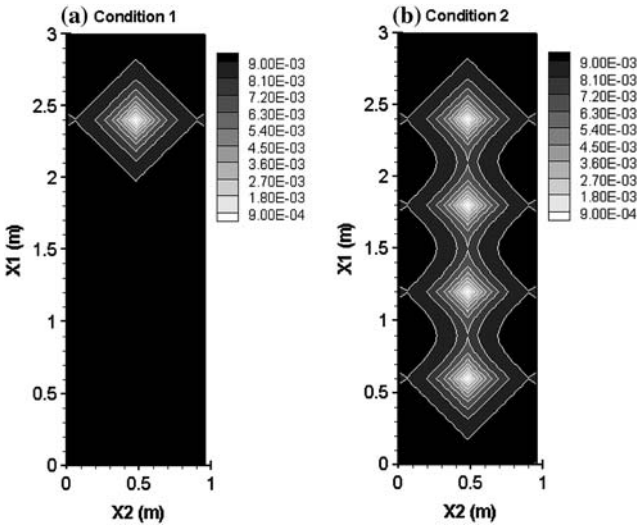

Fig. 5 The variances of log pore size distribution β with **a** 1 measurement, and **b** 4 measurements

still peak at that point. With additional Y conditioned in Case 3, oil pressure variances reach a recess at the oil leak location. In Case 5 and 6, with Y, β and $Y, \bar{n}$ conditioned respectively, the peak of oil pressure variance profile along central vertical line in Case 7 and the recess in Case 3 at the oil leak location are flatten out. These behaviors indicate that the measurement of Y can reduce the oil pressure uncertainty a lot more than that of β or $\bar{n}$ locally, and hence the overall oil pressure uncertainty throughout the modeling domain.

The behavior of capillary pressure variance profile is shown in Fig. 9b. Similar with oil pressure variance, Case 7 produces much larger capillary pressure variance than Case 3, 5, and 6, although it was smaller than unconditional case. In contrast to oil pressure variability, the capillary pressure variability reduces less in Case 5 than in Case 6. Also, the capillary pressure variability reaches the peak at the oil leak location in Case 7. However, the profile of capillary pressure variance in Case 5 and 6 arrives at a recess at the oil leak location, like Case 3. In addition, the variances of capillary pressure in Case 5 and 6 differ from Case 3 less than those of oil pressure in Fig. 9a, indicating the relative importance of uncertainty of intrinsic permeability Y versus the pore size distribution β or the van Genuchten fitting parameter $\bar{n}$ to the uncertainty of capillary pressure is bigger than the oil pressure.

4.3 Correlation lengths of random fields

To examine the influence of correlation lengths of random soil properties, we conduct Baseline Case II and Case 8 with half correlation length of that in Baseline Case and Case 1 (Table 2). Figure 10a shows that the distribution of

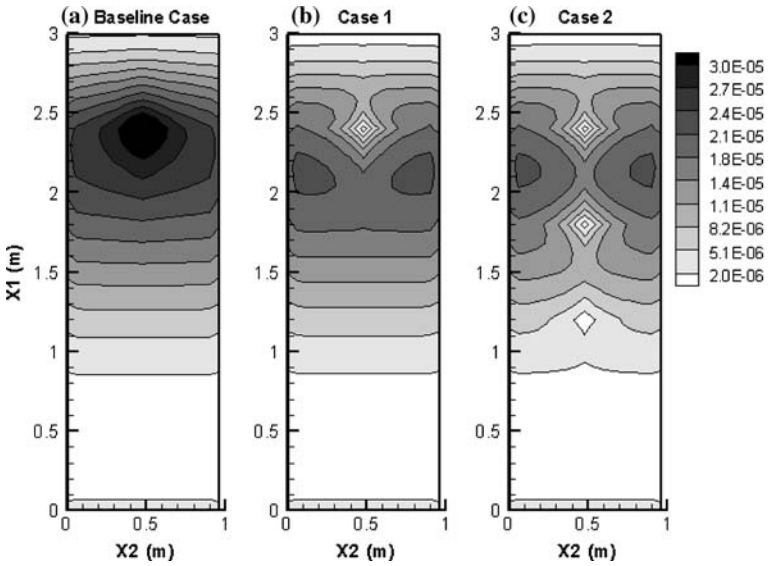

Fig. 6 The contour map of variances of water saturation from **a** unconditional Baseline Case, **b** conditional Case 1 with one measurement, and **c** conditional Case 2 with four measurements

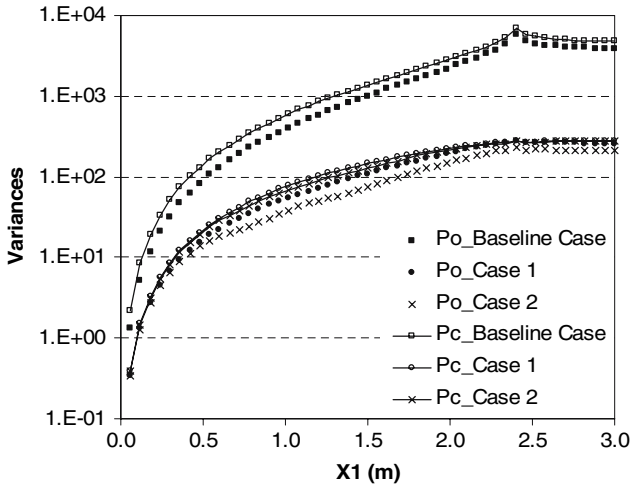

Fig. 7 The variances along central *vertical line* (X2 = 0.48 m) of oil and capillary pressures from unconditional Baseline Case, conditional Case 1 and Case 2

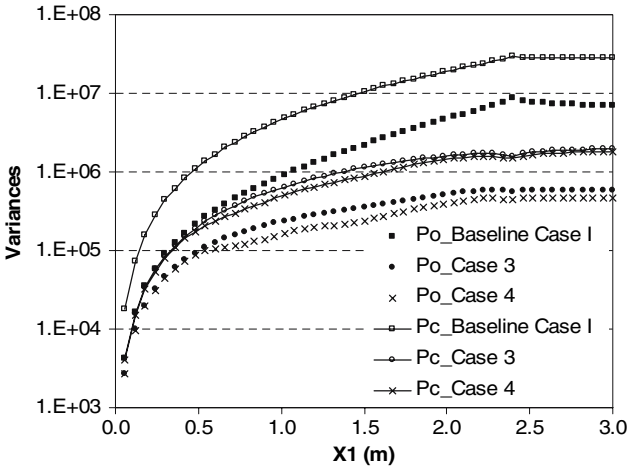

Fig. 8 The comparison of variances along central *vertical line* (X2 = 0.48 m) of oil and capillary pressures between unconditional Baseline Case I, conditional Case 3 and Case 4

variances of water saturation for Baseline Case II is radial centered at oil leak location, similar to that for Baseline Case (Fig. 6a). Oil saturation variance can be proved to be the same as the water saturation variances in such a water–oil flow system. The magnitude of variance in Baseline Case II is slightly smaller than that in Baseline Case, since random fields with smaller correlation length reduce the prediction variability (Zhang and Lu 2004). Another possible reason is because more terms are needed in KL to capture the variability of random fields with shorter correlation lengths. In Baseline Case II, we use the same number of terms (200) as that in Baseline Case, so the final assembles of variances are a little smaller. Similarly, the variability of water saturation from Case 8 (Fig. 10b) is a little smaller than that from Case 1 (Fig. 6b). Owing to the measurements of the three of soil random fields at the oil leak location, the variances fall substantially but very localized around the conditioning points.

What should be noted in the two figures is that the influencing region of the measurement at oil leak location in Case 1 is larger than that in Case 8, which is expected, since the larger correlation length indicated the larger affecting area of each conditioning point, which would reduce relatively the overall prediction variability more. The average relative reduction of oil and capillary pressure variability profile along central vertical line of conditional Case 1 from unconditional Baseline Case are 92.50 and 93.21% (Fig. 7). The reductions of Case 8 from Baseline Case II are 91.29 and 90.58% (Fig 11), which are smaller than the counterpart with larger correlation length (Case 1 vs. Baseline Case). The peaks of oil pressure and capillary

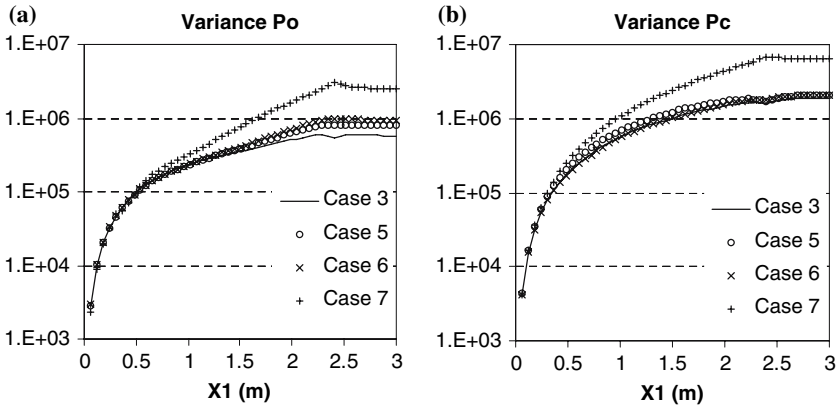

Fig. 9 The comparison of variances along central *vertical line* (X2 = 0.48 m) between Case 4, Case 5, Case 6, and Case 7 for **a** oil pressure, and **b** capillary pressure

pressure profiles at the oil leak location in Case 8 is obvious in Fig. 11, while the peaks in Case 1 is almost damped in Fig. 7. This demonstrates that small correlation length of random input fields has small reduction of prediction uncertainty locally, in addition to the small influence area. Comparison between Figs. 7 and 11 indicates that there are few differences of oil and capillary pressure variability behavior other than that mentioned above.

5 Summary and Conclusions

In modeling subsurface multiphase flow, the hydrogeologic properties characteristic of large heterogeneity lead to huge uncertainty of migration of flowing multiple fluids. Reducing the uncertainty and making predictions as accurate as possible are the major concerns for engineers in remediation and reservoir simulation. In this study, we manage to reduce the prediction variability by incorporating existing measurements of log intrinsic permeability, pore size distribution, and van Genuchten parameters into the unconditional KLME method for the water–oil flow system developed earlier. The key idea of the algorithm is to compute the conditional covariance of the random soil properties via a kriging method, and decompose the conditional covariance into the conditional eigenvalues and eigenfunctions. After solving a series of decomposed deterministic multiphase flow equations with these conditional eigenvalues and eigenfunctions, we can assemble prediction mean and variances of dependent variables (water, oil and capillary pressure; water and oil saturation).

In this paper, we apply a conditional KLME method to multiphase flow systems for the first time. The examples are hypothetical and supposed only for proof of concept, and far away from practical problems, although, it provides a potential tool for remediation and petroleum reservoir engineers to better understand the multiphase flow in heterogeneous subsurface environment, especially under such

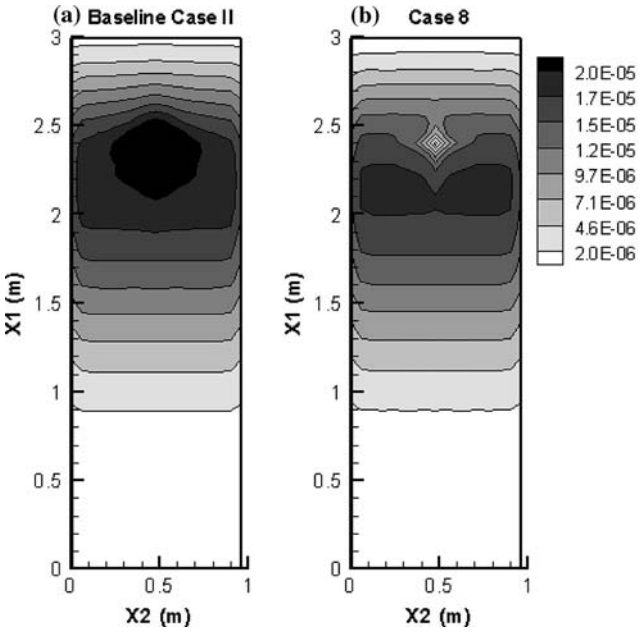

Fig. 10 The contour map of variances of water saturation from **a** unconditional Baseline Case II, and **b** conditional Case 8

a situation that some direct measurements of soil properties are available. A series of simple cases were simulated to investigate how the random soil properties and their measurements influence the stochastic behavior of predictions, which are summarized as follows:

1. The measurements can reduce the prediction variability in the neighboring area remarkably, as well as the overall prediction variability across the domain. More measurements, more reduction of variability. Some measurements in the key location (e.g. source term location) can reduce the uncertainty much more effectively.
2. In our study, the uncertainty of intrinsic permeability seems to contribute more to the uncertainty of prediction than the other two random variables do, while the conditional pore size distribution and van

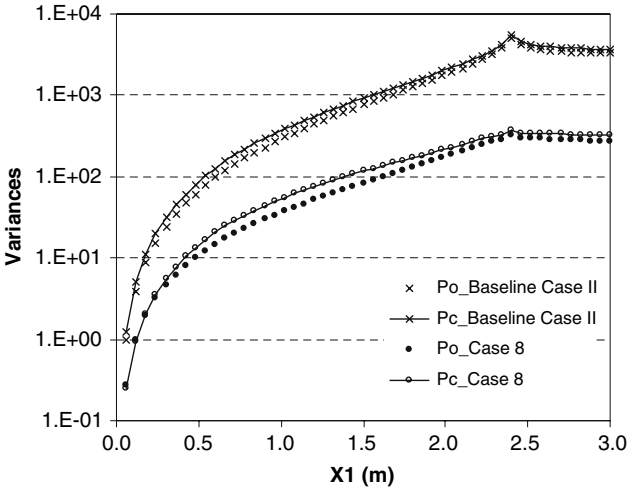

Fig. 11 The variances of oil and capillary pressures along central *vertical line* (X2 = 0.48 m) from unconditional Baseline Case II, and conditional Case 8

Genuchten parameter field only make the slight difference of uncertainty reduction of prediction. Thus, adding measurements to the random intrinsic permeability fields can reduce the prediction variability more than the other two random fields.

3. The increase of source term strength can increase the variances of prediction by several orders, and counteract the uncertainty reduction of conditioned random input variables many times. However, the source of prediction uncertainty is from the random input, the increase of source term strength only amplifies it.
4. Small correlation lengths of random fields can reduce the prediction variability. Also, more terms in KLME approach needed for smaller correlation length of random fields to capture the uncertainty of fluid flow. The smaller the correlation length, the smaller area of one measurement can influence, and the smaller the reduction of the uncertainty.

Acknowledgment The authors would like to acknowledge the funding from the oil shale project cooperated between Chevron Energy Technology Company and Los Alamos National Laboratory.

References

Abriola L (1989) Modeling multiphase migration of organic chemicals in groundwater systems—a review and assessment. Environ Health Perspect 83:117–143

Abriola L, Pinder G (1985) A multiphase approach to the modeling of porous media contamination by organic compounds: 1: equation development. Water Resour Res 21(1):11–18

Chen M, Zhang D, Keller AA, Lu Z (2005) A stochastic analysis of steady state two-phase flow in heterogeneous media. Water Resour Res 41:w01006, doi:10.1029/2004WR003412

Chen M, Keller AA, Zhang D, Lu Z, Zyvoloski GA (2006) A Stochastic analysis of transient two-phase flow in heterogeneous porous media. Water Resour Res 42:W03425, doi:10.1029/2005WR004257

Dagan G (1982) Stochastic modeling of groundwater flow by unconditional and conditional probabilities: 1. Conditional simulation and the direct problem. Water Resour Res 18(4):813–833

Dagan G (1989) Flow and transport in porous formations. Springer, New York

Gelhar W (1993) Stochastic subsurface hydrology. Prentice-Hall, Englewood Cliffs

van Genuchten M (1980) A closed form solution for predicting the hydraulic conductivity of unsaturated soils. Soil Sci Soc Am J 44: 892–898

Ghanem R, Dham S (1998) Stochastic finite element analysis for multiphase flow in heterogeneous porous media. Transport Porous Media 32:239–262

Ghanem R and Spanos D (1991) Stochastic finite elements: a spectral approach. Springer, New York

Guadagnini A, Neuman SP (1999a) Nonlocal and localized analyses of conditional mean steady-state flow in bounded, randomly nonuniform domains: 1. Theory and computational approach. Water Resour Res 35(10):2999–3018

Guadagnini A, Neuman SP (1999b) Nonlocal and localized analyses of conditional mean steady-state flow in bounded, randomly nonuniform domains: 2. Computational examples. Water Resour Res 35(10): 3019–3039

Karhunen K (1947) Uber lineare methoden in der wahrscheinlichkeitsrechnung. Am Acad Sci, Fennicade, Ser. A, I, Vol 37, 3–79, (Translation: RAND Corporation, Santa Monica, California, Rep. T-131, August 1960)

Loeve M (1948) Fonctions aleatorires du second ordre, supplement to P. Levy. Processus Stochastic et Mouvement Brownien, Paris, Gauthier, Villars

Lu Z, Neuman SP, Guadagnini A, and Tartakovsky DM (2002) Conditional moment analysis of steady state unsaturated flow in bounded, randomly heterogeneous soils. Water Resour Res 38(4) doi:10.1029/2001WR000278

Lu, Z and Zhang D (2004) Conditional simulations of flow in randomly heterogeneous porous media using a KL-based moment-equation approach. Adv Water Resour 27:859–874

Tartakovsky DM, Neuman SP, and Lu Z (1999) Conditional stochastic averaging of steady state unsaturated flow by means of Kirchhoff transformation. Water Resour Res 35(3):731–745

Yang J, Zhang D, and Lu Z (2004) Stochastic analysis of saturated–unsaturated flow in heterogeneous media by combing Karhunene–Loeve expansion and perturbation method. J Hydrol 29: 418–38

Zhang D (2002) Stochastic methods for flow in porous media: coping with uncertainties. Academic, San Diego ISBN: 012-7796215

Zhang D, and Lu Z (2004) Evaluation of higher-order moments for saturated flow in randomly heterogeneous media via Karhunen–Loeve decomposition. J Comput Phys 194(2):773–794